

GROUPS REACTIONS TO FAILURE: A SOCIAL IDENTITY APPROACH

by

Noah B. Lambert

BS, University of Pittsburgh, 2000

Submitted to the Graduate Faculty of
Arts and Sciences in partial fulfillment
of the requirements for the degree of
Masters of Science

University of Pittsburgh

2005

UNIVERSITY OF PITTSBURGH

FACULTY OF ARTS AND SCIENCES

This thesis was presented

by

Noah B. Lambert

It was defended on

February 21, 2005

and approved by

John Levine, Professor, Psychology

William Klein, Assistant Professor, Psychology

Thesis Advisor: Richard Moreland, Professor, Psychology

GROUP REACTIONS TO FAILURE: A SOCIAL IDENTITY APPROACH

Noah Lambert, MS

University of Pittsburgh, 2005

This experiment examined how people react after their group has failed at a task. One-hundred thirteen female participants were randomly placed in three-person groups and asked to complete an aesthetic judgment task similar to that used by Klein (1997). All participants then received (false) feedback showing that their group had failed on the task, with many participants also receiving individual feedback about their performance. Some participants were told they performed better than the other members of their group, some were told they did worse, and the rest received no individual feedback. Half of the participants were also told they would lead their group, if it performed the task again, whereas others were told that another member would lead the group. Participants were then asked to choose between trying the same task again, or a taking chance option to win a cash prize. Social identity theory suggests that individuals who “performed” better than other group members, and were also selected as leader, might try to enhance their personal and social identities by choosing to do the same task again. In contrast, individuals who “performed” worse, and were selected as leader, would be less likely to choose to do the task again. Results showed simply that those who performed better were more likely than those who performed worse to try the task again. The leadership variable showed no significant effects.

TABLE OF CONTENTS

Groups reactions to failure: A social identity approach.....	1
Method	6
<i>Participants</i>	6
<i>Materials</i>	6
<i>Procedure</i>	7
Results	10
<i>Response Interdependence</i>	10
<i>Response Distributions</i>	10
<i>Manipulation Checks</i>	11
<i>Responses to Questionnaire 1</i>	11
<i>Responses to Questionnaire 2</i>	13
Discussion	15
<i>Strengths and Weaknesses</i>	17
<i>Future Research</i>	19
APPENDIX A	21
Tables	21
APPENDIX B	24
Figure	24
BIBLIOGRAPHY	25

LIST OF TABLES

Table 1: <i>Overall Means and Standard Deviations for Responses to Questionnaire 1</i>	21
Table 2: <i>Means and Standard Deviations by Condition for Responses to Questionnaire 1</i>	22
Table 3: <i>Overall Means and Standard Deviations for Responses to Questionnaire 2</i>	23
Table 4: <i>Mean and Standard Deviation by Condition for Task Preference Index Scores</i>	23

LIST OF FIGURES

Figure 1: <i>Chart showing the cells of the research design</i>	24
---	----

GROUP REACTIONS TO FAILURE: A SOCIAL IDENTITY APPROACH

Work groups have become an important part of daily life in today's world. Many tasks are now performed by people working in groups. Some tasks can only be performed by groups, and other tasks are performed better by groups than by individuals. As Hackman (1990, 1998) and others have noted, however, work groups can and do fail. This fact has led to considerable research on group effectiveness (see Guzzo, 1986; Hirokawa, 1982; Larson & LaFasto, 2001). A common goal of such research is to learn how to keep groups from failing at their tasks. In contrast, the goal of my research is to learn more about how groups react to task failures.

Reactions to task failures by groups are seldom studied, yet some theory and research can be found. Some of this work (Sherman & Kim, 2005) has focused on impression management tactics by group members (see Snyder, Lassegard, & Ford, 1986; Wann, Hamlet, Wilson, & Hodges 1995), such as cutting off reflected failure (when people try to weaken their association with an unsuccessful group) and group-serving biases (when members attribute group failure to external rather than internal factors). Another line of research involves work groups in declining organizations. Greenhalgh (1983) reviewed the research in this area and found that many groups in such organizations change in ways that cause them to do worse, rather than better. For example, commitment and identification with the organization often decline among workers, which leads to less innovation and more turnover (especially among the best workers). Finally, some research has focused on the "escalation of commitment" that occurs when a course of action chosen by a group has led to losses, but there is still a possibility of success if more time, money, or effort are invested (Dietz-Uhler 1996).

Many of these studies suggest that social identity can play a significant role in how members react to group's failure. Tajfel and Turner (1979, 1986) were among the first to analyze social identity in groups. Their theory was centered around three concepts: social categorization, social identity, and social comparison. Social categorization is a way of perceiving the self and others that emphasizes group memberships rather than individual characteristics. When people engage in social categorization, they become concerned about their social identities. Social identity is that part of the self that reflects the groups to which a person belongs. People want to feel good about themselves (Leary, 1995; Tesser, 1988), so they worry about the quality of their groups. Social comparison of ingroups to relevant outgroups allows people to make quality judgments.

What happens when someone's social identity is threatened because the quality of his or her groups seems poor? According to Tajfel and Turner (1979, 1986), that person can (a) leave those groups for other groups that are better (social mobility), (b) work to make his or her groups better and/or other groups worse (social competition), or (c) change the comparison process itself in various ways (social creativity). Such changes include comparing ingroups to outgroups on new dimensions, redefining comparison dimensions so that ingroups seem superior to outgroups, or comparing ingroups to different outgroups.

How is this theory relevant to task failures by groups? Failing at some task probably threatens the social identities of group members. Social mobility is one way to cope with that threat, but leaving one group to join another is not always an option (see Ellemers, Spears, & Doosje, 1997; Ellemers, Van Knippenberg, De Vries, & Wilke, 1988). When such movement is difficult or impossible, a special kind of social mobility can occur, namely disidentification with the problematic group. A person could de-emphasize membership in that group as a source of social identity, focusing instead on another, better group to which he or she belongs. Or the

person could focus on the quality of his or her personal identity (see Brewer & Weber, 1994). For example, Brown, Collins, and Schmidt (1988) studied self-enhancement tactics among people whose personal identities were good or poor. They found that people with high self-esteem engaged in forms of self enhancement that implicated the self, whereas people with low self-esteem engaged in forms of self enhancement that implicated their group. When individual performance information is available, a person might thus focus on the fact that even though the group failed, he or she succeeded, allowing for a positive personal identity.

Even when individual performance information is not available, research on the self-serving bias suggests that someone in a failing group might still believe that he or she did well. The self-serving bias in groups refers to the tendency of group members to claim personal responsibility for group successes, but deny personal responsibility for group failures (see Forsyth & Schlenker, 1977; Leary & Forsyth, 1987; Schlenker, 1975, Schlenker & Miller, 1977). For example, Schlenker (1975) asked four-person groups to complete an ambiguous task, and then gave each group member (false) performance feedback. This included both objective feedback about the group's performance and relative feedback about how the group compared with other groups that were tested. Everyone was then asked to complete a questionnaire that included measures of perceived responsibility for group performance. Questionnaire responses showed that participants felt more personal responsibility for successful groups and less personal responsibility for groups that failed.

Members of failing groups can also cope with the resulting threat to their social identities in other ways. As Tajfel and Turner (1979, 1986) suggest, social competition is often an option. One socially competitive action is to harm outgroups in some way, so that their performance declines. But such actions are potentially dangerous, because outgroups may retaliate. Most people (see Brewer, 1979) prefer a different socially competitive action, namely helping the

ingroup in some way, so that its performance improves. After failing at a task, for example, a group might decide to try the task again, but work at it harder, or use a better strategy. One such strategy might be to ask someone to lead the group. If a person performed well individually on the task, then group members might want that person to lead the group the second time around. The stronger the leader's own skills, the more likely he or she would be to help the group succeed.

Someone who is good at a group's task should embrace the leadership role, because he or she can help the group to succeed, which would improve that person's social and personal identities. Someone who is poor at the task, however, should avoid the leadership role, because if he or she cannot help the group to succeed, then that person's social and personal identities would worsen.

In summary, group task failure can threaten the social identity of group members. To cope with this threat, members may rely on such tactics as focusing on personal identity or having a leader guide their group. This analysis led me to design an experiment. The procedure for that experiment came in part from research on social comparison by Klein (1997). All participants in my experiment were made to believe that their group had failed at a task, but some participants also received individual performance feedback. Some of these participants were told that they did better than other members of their group, and some were told that they did worse. The remaining participants received no individual performance feedback. Every participant was also told whether or not she would soon become the group's leader. Finally, participants were asked to choose between trying the same task again versus relying on chance to try to win a cash prize.

The experiment thus followed a 3 X 2 design. The first independent variable was Individual Feedback, which was either "better", "none", or "worse." The second independent

variable was Leadership Status--participants either were or were not chosen to lead their groups. The main dependent variable was task choice: evaluating pictures again or selecting poker chips. The probability of succeeding by chance was the same for both of these tasks, but performance on the picture judging task depended on skill, as well as luck.

I expected poor group performance to threaten each participant's social identity. But participants who were told that their personal performance was good might still have a positive personal identity. Participants who were told nothing about their personal performance might also have a positive personal identity, because of the self-serving bias. In contrast, participants who were told that their personal performance was poor should have a negative personal identity.

To cope with the threat to their social identity, participants whose personal performance was good might want to try the task a second time, because they could help the group succeed. This should be especially true when these individuals were selected as group leaders, because that would allow them to help the group (and show off their skills), thereby improving both personal and social identity. The same prediction could be made about participants who received no feedback about their personal performance. But participants whose personal performance was poor should be less eager to try the task again, especially when they were selected as group leaders. They might fear harming the group, and disgracing themselves, causing further damage to their personal and social identities. Thus, poor performers who were not selected as group leaders should be more interested in trying the task again. They know that other members of the group are more skilled, and that one of those members will soon lead the group, probably to success.

I did not predict any statistical main effects in the experiment, but rather an interaction effect (see Figure 1). Leaders in the "better" and "none" individual feedback conditions should show a preference for the picture task option, compared to their non-leader counterparts (Cell A

> Cell B and Cell C > Cell D). In contrast, leaders in the “worse” condition should prefer the poker chip option, compared to non-leaders in the “worse” condition (Cell F > Cell E).

METHOD

Participants

Participants were 113 female undergraduates at the University of Pittsburgh. Participation helped these students to meet an Introduction to Psychology course requirement. The students were randomly assigned to three-person groups, which were then randomly assigned to the cells of the research design.

Materials

Transparencies of the picture pairs of the Meier Art Judgment test, used by Klein (1997) and others, were prepared. These were grayscale drawings of various scenes, such as water flowing under a bridge or a field full of sheep. The two pictures in each pair were the same except for small differences, such as a tree on the left side of a picture instead of the right. An overhead projector was used to display these pictures on a screen to the groups.

Two questionnaires were used in the experiment. The first questionnaire was similar to the one used by Klein (1997). Several questions measured the participants’ satisfaction with their task performance: “How good would you say you are at aesthetic judgments?”; “How good do you think you are at this picture judging task?”; “How happy were you with your group’s performance at this task?”; and “How confident are you in your ability on this task?” Participants were also asked: “How responsible were you personally for the group’s performance on this task?”; “How important was it to you to do well at this picture judging task?”; and “How

related is this task to aesthetic judgment?” Responses to all of these questions were made on seven-point (1 to 7) scales.

The second questionnaire asked which of two task options participants would prefer to attempt to win a prize, and then asked them to rate how confident they were about that choice. Participants were also asked, “Which of these two options do you think is more likely to help your group win the prize?” and “Which option would you enjoy more?” The confidence rating and responses to these last two questions were made on seven-point (1 to 7) scales.

Procedure

Each group was first presented with the 20 picture pairs. Group members were told to record on an answer sheet which picture in each pair was most aesthetically pleasing. Each pair of pictures was presented for 20 seconds (or until the last person recorded her answer), with a few seconds between each presentation. The picture pairs were presented in the same order for every group. All participants were told that there was a correct choice in each case, as determined by a panel of professional artists. They were also told that “group choices” would depend on the pictures that the majority of them preferred, as identified later by the experimenter. All participants expected to receive group performance feedback.

After the last pair of pictures was presented, the experimenter collected the participants’ answer sheets and sat at the end of the table to score them. After three minutes, he completed three (false) feedback sheets, each of which included a group score. In the individual feedback conditions, an individual score was included on each feedback sheet, along with the group score. The feedback sheets also identified the leader of each group by including either the letter “Y” or the letter “N”, indicating whether or not (Yes or No) each person had been randomly chosen by the experimenter to serve as group leader.

Before distributing these feedback sheets, the experimenter told all of the participants that:

“As a group, you left some room for improvement. Your group only got 8 of 20 correct. Individually, one of you got 13, one of you got 7, and one of you got 5 correct. I have a feedback sheet here for each of you; please keep these sheets to yourself.”

The feedback sheets were then distributed to participants. Note that because the group score was determined by the majority preference in each picture pair, it was possible for the group to earn a poor score, even if one group member performed well. Therefore, participants realized that their personal and group scores need not be compatible.

Two manipulations were carried out using the feedback sheets. The first manipulation involved the individual feedback for each group member. Each person learned that she answered correctly either 5 times (worse condition), or 13 times (better condition), or she received no individual feedback. All three members of each group actually received the same feedback, but each person was meant to believe (because the individual feedback information was not shared) that the other members of her group received the other two scores mentioned by the experimenter. The second manipulation involved group leadership. Members of groups in the leader condition all had a “Y” on their feedback sheets. Because they did not share the information on those sheets, each of these participants was meant to believe that she had been selected as the group leader. Members of groups in the no-leader condition all had an “N” on their feedback sheet, so each of these participants was meant to believe that someone else had been selected as the group leader.

Next, participants were given the first questionnaire. They had up to five minutes to complete that questionnaire, and they were allowed to keep their feedback sheets during that period. Afterwards, participants were told that they had an opportunity to win a \$10 (per group

member) prize. As in Klein's (1997) experiment, participants were given a choice between two tasks for which they might receive this prize. Their first option was to complete another 20 aesthetic judgments--the group had to select the best picture at least 10 times to win the prize. By chance, there was thus a 50% probability of succeeding. But the participants were also told that this time they would be allowed to discuss their choices together as a group. They learned that one member of their group had been randomly selected as the group's leader, someone who would be responsible for guiding the group's discussion towards the correct choices. They were told that the letter (Y or N) on each feedback sheet indicated whether or not that person would be the leader of the group.

The second option for winning the prize was to simply pull blue chips out of a bag containing an equal number of blue and red chips. Someone in the group had to pull a blue chip out of the bag at least 10 times out of 20 in order to win the prize. Because there was always an equal number of red and blue chips in the bag, the group again had a 50% probability of succeeding, but performance on this task depended entirely on luck.

After these options were explained to them, group members were given the second questionnaire and had up to five minutes to complete it.

After the second questionnaire was collected by the experimenter, all participants were debriefed, thanked for their help, and dismissed. No one seemed suspicious about the procedures, nor could anyone guess any research hypothesis. Because they did not have the opportunity to actually win a prize, participants' names were entered into a lottery. In that lottery, five winners were drawn, and each of them was mailed a \$20 payment.

RESULTS

After examining responses to the questionnaire items, it became apparent that one participant did not take the experiment seriously. She was thus dropped from the sample, leaving a final sample of 113 students.

Response Interdependence

Although group members did not interact with one another, there may have been some interdependence in their responses. If so, then that would have to be taken into account in later data analyses. Following methods described by Kenny and Judd (1986; see also Kenny, Mannetti, Pierro, Livi, and Kashy, 2002) a one-way ANOVA was performed on each dependent measure, using group membership as the independent variable. The mean square within groups was multiplied by the number of people in each group minus one (3-1), then added to the mean square between groups value. This total was divided into the mean square between groups minus the mean square within groups to calculate an intraclass correlation for each dependent measure. These correlations were small (-0.10 to 0.18), and only one of them was significant. Apparently there was little response interdependence, so the data were analyzed at the individual level.

Response Distributions

The individual responses to each questionnaire item were tested for skewness. Only two items, both on Questionnaire 2, showed significant levels of skewness. The first item asked which option a participant would choose to win the prize, and the second item asked which option was more likely to help a group win a prize. Both items involved binary response formats, which may explain why their response distributions were skewed. Research has shown that ANOVAs can still produce sound results when the dependent measure is dichotomous,

especially when the degrees of freedom for the within cell variance is equal to 20 or more (Lunney, 1970). The data from my experiment met that criterion

Manipulation Checks

Item 5 on Questionnaire 2 asked each participant how her individual score on the picture judging task compared with the scores for the other members of her group. Every participant in the “better” condition said that her score was better than the scores for the other members of her group, and every participant in the “worse” condition said that her score was worse. Among the participants who received no individual performance feedback, 62% (24 of 39) said that no information was available to answer this question. One participant said her score was lower than the scores for the other members of her group. The remaining participants in this condition said that their scores fell between those of the other group members.

Item 6 on Questionnaire 2 asked participants whether they, or someone else in their group, had been chosen as the leader. Every participant in the “no leader” condition said that she was not chosen as group leader. Among the participants in the “leader” condition, 95% (53 of 56) said that they were chosen as group leader. This manipulation thus seemed successful as well.

Later analyses were performed on data from varying samples that excluded some or all of the participants also answered the manipulation checks incorrectly. Similar results were found for all of these samples, so only results for the largest sample (N=113) are reported here.

Responses to Questionnaire 1

The overall means and standard deviations for responses to Questionnaire 1 can be seen in Table 1. In general, participants did not feel very skilled and were not confident about their

task ability. T-tests comparing the means of the first three items to the midpoint (4) of the rating scale showed that these means were all significantly below that point [$t(112) = -3.36, p < .01$; $t(112) = -6.04, p < .01$; and $t(112) = -8.30, p < .05$ respectively]. The mean for the happiness measure was also significantly below the scale's midpoint $t(112) = -16.12, p < .01$, showing that participants were generally unhappy with the performance of their groups. This suggests that they heard and believed that group performance was poor. But as the self-serving bias might suggest, participants did not feel very responsible for the performance of their groups. The overall mean for this measure was below the scale midpoint, though not to a significant degree. Motivation among participants was also low; in general, they did not feel that it was very important to do well on the task. The mean for this measure was significantly lower than the scale midpoint ($t(112) = -4.78, p < .01$). This lack of motivation may have been related to participants' beliefs about the task. As the mean overall response on the last questionnaire measure suggests, many participants were not convinced that the task was related to aesthetic judgment.

The means and standard deviations by condition for responses to Questionnaire 1 are shown in Table 2. [Note that at this point in the experiment, the leader manipulation had not yet occurred]. Each measure was analyzed using a single factor ANOVA, with Individual Feedback (worse, none, better) as the independent variable. These analyses revealed significant differences on skill at aesthetic judgment, $F(2, 110) = 20.63, p < .01$; skill at the picture judging task, $F(2, 110) = 29.57, p < .01$; confidence in ability, $F(2, 110) = 7.08, p < .01$; and beliefs about whether the task was related to aesthetic judgment, $F(2, 110) = 7.18, p < .01$. Tukey post hoc tests showed that on the first measure (skill at aesthetic judgment), the mean for the "worse" condition was significantly ($p < .01$) lower than the means for the "none" and "better" conditions, which did not differ significantly ($p > .05$) from one another. On the second measure (skill at the picture

judging task), the means for all three conditions (worse, none, better) differed significantly from one another ($p < .01$), with “worse” showing the lowest mean and “better” the highest mean. On the third measure (confidence in ability), the mean for the worse condition was significantly lower than the mean for the “better” condition ($p < .01$); the mean for the “none” condition was not significantly different from either of those means. Finally, on the fourth measure (task related to aesthetic judgment), the means for the “worse” and “none” conditions were significantly lower than the mean for the “better” condition ($p < .01$), but did not differ significantly from one another.

Note that the results for the skill at aesthetic judgment, skill at the picture judging task, and confidence in ability measures all provided additional support for the effectiveness of the individual feedback manipulation.

Responses to Questionnaire 2

Item 7 on Questionnaire 2 was used to check whether participants heard and believed that they could discuss their aesthetic judgments as a group if they chose to perform the task a second time. Nearly all of them (104 of 113, or 92%) answered this question correctly.

Table 3 shows the overall means and standard deviations for responses to the dependent measures (Items 1-4) on Questionnaire 2. Task choice was assigned a value of “0” if the poker chips option was chosen, or “1” if the picture task option was chosen. Participants generally preferred the picture task option (77 vs. 36; $\chi^2(1) = 14.88, p < .01$), although they were not very confident about their choices. A t-test comparing their mean confidence score to the midpoint (4) of the rating scale showed that the mean was significantly below that point, $t(112) = -2.04, p < .05$, suggesting a lack of confidence. Participants also preferred the picture task option when asked which option would be more likely to win them the prize. The mean response on this

measure was significantly above the midpoint of the scale, $t(112) = 4.71, p < .01$. However, when participants chose which task they would enjoy more, there appeared to be no preference. The mean on this measure (3.84) was not significantly different ($p > .05$) from the scale midpoint. Nevertheless, these last two items were significantly correlated ($r = .34, p < .05$) suggesting that enjoyment of a task was related to the perceived likelihood that it would produce a prize.

All four items were combined into a single, task preference index. This process involved first assigning task choice a value of “-1” for the poker chips and “+1” for the picture task. The confidence scores were then reduced by one, changing the old range of 1 to 7 to a new range of 0 to 6. This allowed for a score of 0 if a participant had no confidence in her choice. The new task choice values were then multiplied by the new confidence values to create scores that incorporated both the participants’ choices and their confidence in those choices. These new scores, along with the scores on the other two items, were all standardized, and then the standardized scores were averaged to create an index ($\alpha = .70$). Higher scores on this index showed a greater preference for the picture judging task.

The means and standard deviations for index scores across conditions are shown in Table 4. A 2 X 3 (Leadership X Individual Feedback) ANOVA on these scores revealed only a main effect for Individual Feedback, $F(2, 107) = 5.48, p < .01$. A Tukey post-hoc test showed that the mean for the “worse” condition was significantly lower than the mean for the “better” condition ($p < .01$). The mean for the “none” condition was not significantly different from either of those two means.

The same analysis, using simpler scores that were based only on task choice and confidence as the dependent measure produced very similar results.

Scores on the task choice index correlated significantly with scores on the task skills index mentioned earlier, $r = .37, p < .01$. This suggests that participants who believed they were skilled at judging pictures were likely to choose the pictures option the second time around. The task choice index scores were also significantly correlated with participants' confidence in their task ability, $r = .27, p < .01$. This shows that participants with greater confidence were more likely to choose the pictures option as well. Finally, beliefs about how related the picture task was to aesthetic judgment were significantly correlated with responses to how important it was to do well on that task, $r = .29, p < .01$. The more that participants thought the picture task was related to aesthetic judgment, the more they felt it was important to do well at that task. A few other significant correlations were also found, but they were not of much interest.

DISCUSSION

Participants in this experiment were asked to perform an aesthetic judgment task as members of small groups. After completing that task, group members were given feedback indicating that their groups had failed. Participants were affected by that failure. Analyses of questionnaire responses showed that participants' ratings of their task skills and confidence were generally low (significantly below the midpoint of the scale). Moreover, participants were generally unhappy with the performance of their groups.

Some participants were also given feedback about their personal task performance. Manipulation checks showed that this feedback also had an impact. All participants in the "better" condition said that they performed better than the rest of their groups, and all participants in the "worse" condition said they performed worse than the rest of their groups. Among participants who received no feedback, most said that no individual information was

available, or that their performance was average for their groups. I also combined participants' task skills responses to create an index showing how good participants believed they were at the task. An analysis of the index scores showed that participants in the "better" condition viewed their task ability most positively, followed by participants in the "none" condition, and then by participants in the "worse" condition. Again, this suggests that the manipulation was successful.

Participants were also told whether they would be group leaders or not. This manipulation seemed to be successful as well—all participants in the "no leader" condition said that they were not chosen as group leaders, and all but three participants in the "leader" condition said that they were chosen as group leaders.

An unexpected result was that participants did not vary significantly across the different conditions in how responsible they felt for the performance of their groups. The self-serving bias was apparently stronger than I expected—even participants in the "worse" condition, who had reason to feel responsible for the poor performance of their groups, did not accept much responsibility.

The dependent variable was the task option that participants chose to seek the prize. Participants who chose to do the pictures task again probably felt confident about the ability of their groups to do well. In contrast, participants who chose the poker chips option were probably skeptical about the ability of their groups. Overall, the pictures option was more popular than the poker chips option. Several questionnaire items were combined to create an index representing task choice and confidence in that choice. An analysis of index scores showed an unexpected main effect of individual feedback on task choices. Participants who were told that they performed worse than other group members had less desire to do the pictures task again, whereas participants who were told that they performed better than other group members had more desire to do that task again. Participants who received no feedback fell between the other feedback

conditions. Although this finding was unexpected, it seems consistent with my hypothesis. I reasoned that people who felt they performed worse than other members of their group would be less likely to want to try the task again, whereas people who felt that they performed better would be more likely to want to try the task again. This is in fact what happened, although leadership did not moderate the effect.

The predicted interaction was not confirmed by the results. Participants in the “worse” individual feedback condition were expected to show less desire to do the pictures task again when they were chosen as group leaders than when they were not. The opposite pattern actually occurred, although the difference was not significant. Conversely, participants in the “better” and “none” individual feedback conditions were expected to show more desire to do the pictures task again when they were chosen as group leaders than when they were not. This pattern of results did occur, although again, the differences were not significant.

Strengths and Weaknesses

A major goal of this experiment was to move research on group performance in a new direction. Most research on poor group performance explores why groups fail (see Guzzo, 1986; Hirokawa, 1982; Larson & LaFasto, 2001). That is an important topic, but it seems just as important to understand how groups react to failure. Every group fails occasionally, so understanding how groups react to failure might help people to avoid some of its negative consequences. Do group members lose confidence in their personal abilities, or the ability of their groups, or both? Does a failing group need different tactics to succeed? Does it need new members? If no changes are made, and the group simply tries harder, then it may well continue to fail (Dietz-Uhler, 1996; Seibert & Goltz, 2001; Whyte, 1993). These are issues that could be resolved by studying how groups react to failure.

One strength of my experiment was the use of a cash prize to motivate participants to perform well. Another strength was having participants make private choices during the pictures task. This made the individual performance feedback believable because group members could not directly compare the responses of others in their group to their own responses. The novelty and ambiguity of the task was also helpful. Participants could not readily tell whether their choices were correct, so I was able to make them believe that they had failed personally and/or as a group. And this feedback was given privately, so that each group member could receive the same feedback as other members of her group, without realizing that fact.

Another procedural strength was how reactions to group failure were measured. Instead of just asking how confident group members were about their own task ability, I also asked them to choose a task option that they felt would help them win a prize. This allowed me to measure how confident participants felt about their own ability and the ability of their groups, without revealing what I was measuring. If participants had realized what I was measuring, then demand characteristics might have affected their responses. Ensuring that questionnaire responses were confidential probably also made for more honest responding, as well as limiting influence among group members.

A procedural weakness of this experiment was that group members did not interact in the ways that people do in more natural groups. And levels of social identity were probably weak overall, because the groups were short-lived and not very meaningful to their members. I could have tried to measure levels of social identity, but that might have revealed the nature of the experiment to participants, which would have affected their responses. Another trade-off was the use of the aesthetic judgment task. Although it made negative feedback believable, the task may not have seemed very relevant to participants (indeed, few of them said that it was important for them to perform well). The relevance of a task could play a key role in how the

personal and social identities of group members are affected. If a task seems irrelevant, then poor task performance is unlikely to threaten either personal or social identities. A related problem involves the poker chips option, which may have seemed less enjoyable than the picture judging task. Pre-testing might have revealed a chance task option that would have been more attractive to participants. Pre-testing might also have helped me choose more appropriate individual performance scores to assign in the “better” and “worse” feedback conditions. The scores were not equally different from each other ($13-7=5$ and $7-5=2$), making it difficult to compare the effects of individual feedback across conditions.

Future Research

Some obvious directions for future research involve solving some of the problems just mentioned. For example, I could focus on trying to create stronger social identity among group members, or asking groups to perform a more relevant task.

One way to solve the problem of weak social identity would be to study pre-existing groups, such as students who work together as part of a class or employees who work together in a company. These groups could be made to feel as if they failed at a relevant task. With a group of students, for example, this might be accomplished by asking them to complete an assignment similar to one they have done in class, and then telling them that they did not do very well on my assignment. Each group member could then be asked to describe how he or she would proceed on a new assignment. Group member responses could be coded as to whether participants want to try the new assignment with similar methods, try it with different methods, or perhaps not try it at all.

If pre-existing groups are not available, then I could try to strengthen social identity in ad hoc groups like those used in my experiment. Dietz-Uhler (1996) showed that this can be

accomplished by labeling group members as a group. She found that even such small acts as not introducing group members to one another, giving group members identical name tags containing just the name of their group, and referring to the group as “your group,” can produce higher levels of social identity.

One puzzling aspect of my experiment was that leadership seemed to have no effects. A closer examination of leader effects would thus be another interesting direction for future research. For example, I could study (pre-existing) groups that already have leaders. The groups could then be made to feel that they failed at a task while working under the direction of those leaders. If they were asked to complete another, similar task, would group members stick with their leaders, or try different leaders? A similar experiment could involve manipulating how leaders arise in groups. Conditions might include an experimenter appointing the leader, group members electing the leader, and a no-leader control condition. After failing at a task, would groups who elected their leader be less likely to get rid of that person? Groups who elected a leader would have more investment in that person and thus might not want to get rid of him or her, whereas groups whose leader was appointed might feel comfortable getting rid of that person. (Hollander, Fallon, & Edwards, 1977)

The effects of leadership could also be studied by manipulating why a group member was selected as leader. Participants could be told that they were selected as group leaders because of their personal task performance. In my study, individual performance seemed to influence the decisions of group members, but being a leader did not. If leadership selection were more closely associated with performance, then it might make being a leader more salient. This could cause leaders to behave as I predicted—“worse” leaders would be afraid of failing in front of their groups, whereas “better” leaders would be more motivated to lead their groups to success.

APPENDIX A

Tables

Table 1: *Overall Means and Standard Deviations for Responses to Questionnaire 1.*

	Mean	S. D.
Skill at aesthetic judgment	3.59	1.29
Skill at picture judging	3.33	1.18
Confidence in ability	2.96	1.34
Happiness with group's performance	2.47	1.01
Responsibility for group's performance	3.84	1.29
Importance of doing well	3.39	1.36
Task related to aesthetic judgment	4.10	1.38

Note: All responses were made on seven-point (1 to 7) rating scales

Table 2: *Means and Standard Deviations by Condition for Responses to Questionnaire 1.*

	<u>"Worse"</u> <u>Condition</u>		<u>"None"</u> <u>Condition</u>		<u>"Better"</u> <u>Condition</u>	
	Mean	S. D.	Mean	S.D.	Mean	S.D.
Skill at aesthetic judgment	2.71	1.23	3.77	1.20	4.33	0.83
Skill at picture judging	2.50	1.03	3.31	1.03	4.22	0.80
Happiness with group's performance	2.42	0.95	2.38	1.14	2.61	0.93
Confidence in ability	2.45	1.31	2.92	1.38	3.56	1.08
Responsibility for group's performance	3.79	1.66	3.72	0.83	4.03	1.25
Importance of doing well	3.32	1.38	3.31	1.36	3.56	1.36
Task related to aesthetic judgment	3.79	1.26	3.77	1.51	4.78	1.10

Note: All responses were made on seven-point (1 to 7) rating scales

Table 3: *Overall Means and Standard Deviations for Responses to Questionnaire 2.*

	Mean	S. D.
Task choice	0.68	0.47
Confidence in choice	3.76	1.25
More likely to win prize	4.64	1.44
Which would enjoy more	3.84	1.92

Note: All responses were made on seven-point (1 to 7) rating scales

Table 4: *Mean and Standard Deviation by Condition for Task Preference Index Scores*

	<u>"Worse" Condition</u>		<u>"None" Condition</u>		<u>"Better" Condition</u>	
Leader	Mean	S. D.	Mean	S. D.	Mean	S. D.
Yes	-0.29	0.54	0.16	0.75	0.35	0.85
No	-0.34	1.06	-0.04	0.73	0.21	0.71

APPENDIX B

Figure

Figure 1: *Chart showing the cells of the research design.*

		Leader?	
		Yes	No
Individual Feedback	Better	A	B
	None	C	D
	Worse	E	F

BIBLIOGRAPHY

- Brewer, M.B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin*, 86, 307-324.
- Brewer, M.B., & Weber, J.G. (1994) Self-evaluation effects of interpersonal versus intergroup social comparison. *Journal of Personality and Social Psychology*, 66, 268-275.
- Brown, J. D., Collins, R. L., & Schmidt, G. W. (1988). Self-esteem and direct versus indirect forms of self-enhancement. *Journal of Personality and Social Psychology*, 55, 445-453.
- Dietz-Uhler, B. (1996). The escalation of commitment in political decision-making groups: a social identity approach. *European Journal of Social Psychology*, 26, 611-629.
- Ellemers, N., Spears, R., & Doosje, B. (1997). Sticking together or falling apart: In-group identification as a psychological determinant of group commitment versus individual mobility. *Journal of Personality and Social Psychology*, 72, 617-626.
- Ellemers, N., Van Knippenberg, A., De Vries, N., & Wilke, H. (1988). Social identification and permeability of group boundaries. *European Journal of Social Psychology*, 18, 497-513.
- Forsyth, D.R. & Schlenker, B.R. (1977). Attributing the causes of group performance: Effects of performance quality, task importance, and future testing. *Journal of Personality*, 45, 220-236.
- Greenhalgh, L. (1983). Organizational decline. In S. B. Bacharach (Ed.), *Research in the sociology of organizations* (Vol. 2, pp. 231-276). Greenwich, CT: Jai Press.

- Guzzo, R.A. (1986). Group decision making and group effectiveness in organizations. In P.S. Goodman & Associates (Eds.), *Designing effective work groups* (pp.34-71). San Francisco: Jossey-Bass.
- Hackman, J.R. (1998). Why teams don't work. In R.S. Tindale & Associates (Eds.), *Theory and research on small groups* (pp. 245-267). New York: Plenum Press.
- Hackman, J.R. (Ed.). (1990). *Groups that work (and those that don't)*. San Francisco: Jossey-Bass.
- Hirokawa, R.Y. (1982). Group communication and problem-solving effectiveness I: A critical review of inconsistent findings. *Communication Studies*, 30, 134-141.
- Hollander, E.P., Fallon, B.J., & Edwards, M.T. (1977). Some aspects of influence and acceptability for appointed and elected group leaders. *Journal of Psychology*, 95, 289-296.
- Kenny, D.A., & Judd, C.M. (1986). Consequences of violating the independence assumption in analysis of variance. *Psychological Bulletin*, 99, 422-431.
- Kenny, D.A., Mannetti, L., Pierro, A., Livi, S., & Kashy, D.A. (2002). The statistical analysis of data from small groups. *Journal of Personality and Social Psychology*, 83, 126-137.
- Klein, W.M. (1997). Objective standards are not enough: Affective, self-evaluative, and behavioral responses to social comparison information. *Journal of Personality and Social Psychology*, 72, 763-774.
- Larson, C. E., & LaFasto, F.M.J. (1989). *Teamwork: What must go right/what can go wrong*. Newbury Park, CA: Sage.
- Leary, M. R., & Forsyth, D.R. (1987). Attributions of responsibility for collective endeavors. In C. Hendrick (Ed.), *Group processes* (pp. 167-188). Thousand Oaks, CA: Sage Publications, Inc.

- Leary, M.R. (1995). *Self-presentation: Impression management and interpersonal behavior*.
Madison, WI: Brown & Benchmark
- Lunney, G.H. (1970). Using analysis of variance with a dichotomous dependent variable: An empirical study. *Journal of Educational Measurement*, 7, 263-269.
- Schlenker, B.R. (1975). Group members' attributions of responsibility for prior group performance. *Representative Research in Social Psychology*, 6, 96-108.
- Schlenker, B.R. & Miller, R.S. (1977). Egocentrism in groups: Self-serving biases or logical information processing? *Journal of Personality and Social Psychology*, 35, 755-764.
- Seibert, S.E., & Goltz, S.M. (2001). Comparison of allocations by individuals and interacting groups in an escalation of commitment situation. *Journal of Applied Social Psychology*, 31, 134-156.
- Sherman, D.K., & Kim, H.S. (2005). Is there an "I" in "Team"? The role of the self in group-serving judgments. *Journal of Personality and Social Psychology*, 87, 108-120.
- Snyder, C.R., Lassegard, M., & Ford, C. (1986). Distancing after group success and failure: Basking in reflected glory and cutting off reflected failure. *Journal of Personality and Social Psychology*, 51, 382-388.
- Tajfel, H., & Turner, J.C. (1979). An integrative theory of intergroup conflict. In W.G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33-48). Chicago: Nelson-Hall.
- Tajfel, H., & Turner, J.C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W.G. Austin (Eds.), *The social psychology of intergroup relations* (2nd ed., pp. 7-24). Monterey, CA: Brooks-Cole.

- Tesser, A. (1986). Some effect of self-evaluation maintenance on cognition and action. In R.M. Sorrentino & E.T. Higgins (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (pp. 435-464). New York: Guilford.
- Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 181-227). San Diego, CA: Academic Press.
- Wann, D.L., Hamlet, M.A., Wilson, T.M., & Hodges, J.A. (1995). Basking in reflected glory, cutting off reflected failure, and cutting off future failure: The importance of group identification. *Social Behavior and Personality*, 23, 377-388.
- Whyte, G. (1993). Escalating commitment in individual and group decision making: A prospect theory approach. *Organizational Behavior and Human Decision Processes*, 54, 430-455.